Г. В. М а р т ы н о в (Москва, ИППИ РАН). **Критерии согласия для распределений Вейбулла–Гнеденко и Парето.**

Рассмотрим выборку $X^n = \{X_1, X_2, \ldots, X_n\}$, состоящую из независимых наблюдений случайной величины с функцией распределения $F(x)$, $x \in \mathbf{R}_1$. Будем проверять гипотезу о принадлежности этого распределения параметрическому семейству распределений H_0: $F(x) \in \mathcal{G} = \{G(x,\theta),\ \theta = (\theta_1, \theta_2, \ldots, \theta_k)^T \in \Theta \subset \mathbf{R}_k\}$ против альтернативы, содержащей все остальные распределения. Для этого рассмотрим статистику омега-квадрат (Крамера–Мизеса) $\omega_n^2 = n\int_{-\infty}^{\infty}(F_n(x) - G(x,\theta_n))^2\,dG(x,\theta_n)$, где θ_n — оценка максимального правдоподобия θ, а $F_n(x)$ — эмпирическая функция распределения. Предельное распределение статистики ω_n^2 зависит в общем случае от неизвестного значения θ_0 параметра θ и семейства распределений $\mathcal{G}$. В 1955 г. Кац, Кифер и Волфовиц [2] установили, что предельное распределение статистики ω_n^2 для семейства вида $\mathcal{Q} = \{Q((x-c)/d),\ -\infty < x < \infty,\ d > 0\}$ не зависит от неизвестных параметров c и d. В частности, таким семейством является множество всех нормальных распределений.

Здесь предлагается другой класс: $\mathcal{R} = \{R((x/\beta)^\alpha),\ \alpha > 0,\ \beta > 0,\ x \in \mathcal{X} \subset [0,\infty)\}$, где $\mathcal{X}$ — носитель распределения $R((x/\beta)^\alpha)$, $R(z)$ является функцией распределения с носителем $\mathcal{Z} \subset [0,\infty)$. Пусть $r(z) = R'(z)$. Критерии Крамера–Мизеса и Колмогорова–Смирнова основаны на использовании эмпирического процесса $\xi_n(x) = \sqrt{n}(F_n(x) - R((x/\widehat{\beta})^{\widehat{\alpha}})))$, где $\widehat{\alpha}$ и $\widehat{\beta}$ являются МП-оценками неизвестных параметров α и β. Ковариационная функция предельного гауссовского процесса $\xi(t)$ есть $K(t,\tau) = \min\{t,\tau\} - t\tau - (1/(B_{11}B_{22} - B_{12}^2))(B_{22}s_1(t)s_1(\tau) - B_{12}(s_1(t)s_2(\tau) + s_2(t)s_1(\tau)) + B_{11}s_2(t)s_2(\tau))$, $t,\tau \in (0,1)$, где $B_{11} = \int_{\mathcal{Z}}(p(z)\log z + \log z + 1)^2 r(z)\,dz$, $B_{22} = \int_{\mathcal{Z}}(p(z)+1)^2 r(z)\,dz$, $B_{12} = \int_{\mathcal{Z}}(p(z)\log z + \log z + 1)(p(z)+1)r(z)\,dz$, $p(z) = zr'(z)/r(z)$, $s_1(t) = r(R^{-1}(t))R^{-1}(t)\log R^{-1}(t)$, $s_2(t) = r(R^{-1}(t))R^{-1}(t)$. Предельное распределение статистики, основанной на $\xi_n(x)$, не зависит от неизвестных параметров α и β.

Результат применим к распределению Парето $F(x) = 1 - (x/\beta)^{-\alpha}$, $x \geqslant \beta \geqslant 0$, $\alpha > 0$, $R(z) = 1 - 1/z$, $\mathcal{Z} = [\beta,\infty]$, к двухпараметрическому распределению Вейбулла–Гнеденко $F(x) = 1 - e^{-(x/\beta)^{-\alpha}}$, $x \geqslant 0$, $\beta \geqslant 0$, $\alpha > 0$, $R(z) = 1 - e^{-z}$, $\mathcal{Z} = [0,\infty]$, к степенному распределению $F(x) = x^\alpha$, $x \in [0,1]$, $R(z) = z$, $\mathcal{Z} = [0,1]$, и т. д. Независимость статистики омега-квадрат от параметров для семейства распределений Парето отмечена в [4].

Существует связь между семействами $\mathcal{G}$ и $\mathcal{R}$. Пусть X является случайной величиной с распределением $R((z/\beta)^\alpha)$. Преобразуем X к случайной величине W с помощью преобразования $W = -\log X$. Тогда $\mathbf{P}\{W < x\} = G((x-m)/\sigma)$, где $G(x) = 1 - R(e^{-x})$ и соответствующие параметры семейства $\mathcal{G}$ есть $m = -\log\beta$, $\sigma = 1/\alpha$. Такое преобразование для распределения Вейбулла было рассмотрено в [1], [5]. Обратное преобразование из $\mathcal{G}$ в $\mathcal{R}$ есть $X = e^W$. Например, семейство нормальных распределений преобразуется в семейство репараметризованных логнормальных распределений. В общем случае также могут использоваться преобразования $W = \log X$ и $X = e^{-W}$.

Работа поддержана РФФИ, проект № 09–01–00740а.

СПИСОК ЛИТЕРАТУРЫ

1. *Choulakian V.*, *Stephens M. A.* Goodness-of-fit tests for the generalized Pareto distribution. — Technometrics, 2001, v. 43, p. 478–484.
2. *Kac M.*, *Kiefer J.*, *Wolfowitz J.* On tests of normality and other tests of goodness-of-fit based on distance methods. — Ann. Math. Statist., 1955, v. 30, p. 420–447.
3. *Мартынов Г. В.* Критерий омега-квадрат. М.: Наука, 1979, 80 с.
4. *Stephens M. A.* Tests based on EDF statistics. — In: Goodness-of-fit Techniques. /Ed. by R. B. D'Agostino, M. Stephens. N.Y.: Marcel Dekker, 1986, p. 97–193.

5. *Тюрин Ю. Н., Саввушкина Н. Е.* Критерии согласия для распределения Вейбулла–Гнеденко. — Изв. АН СССР. Сер. Техн. киберн.,1984, № 3, с. 109–112.